PART I

MEASURING PERFORMANCE AND INTELLIGENCE OF SYSTEMS WITH AUTONOMY

The White Paper

Measuring Performance and Intelligence of Systems with Autonomy: Metrics for Intelligence of Constructed Systems[1]

A White Paper Explaining Goals of the Workshop

1. Introduction

Thousands of person-years have been devoted to research and development in the various aspects of artificially intelligent systems. There is no single field of study that contributes to the progress, but rather several dozens, ranging from control to cognitive sciences. Much progress has been attained. However, there has been no means of evaluating the progress of the field. How can we assess the current state of the science? Some systems are beginning to be deployed commercially. How can a commercial buyer evaluate the advantages and disadvantages of the *intelligent* candidates and decide which system will perform best for their application? If constructing a system from existing components, how does one select the one that is most appropriate within the desired system?

The ability to measure the capabilities of intelligent systems or components is more than an exercise in satisfying intellectual or philosophical curiosity. Without measurements and subsequent quantitative evaluation, it is difficult to gauge progress.

It can even be argued that researchers and developers perpetually re-invent the same components to build their system, unable to reliably find existing components they could reuse. To paraphrase William, Lord Kelvin: when you can measure something and put some numbers to it, then you know something about it, and if you can't your understanding of it is of a "meager and unsatisfactory kind," although I am not sure that I would be so adamant about the need for numbers.

It is both in a spirit of scientific enquiry and for pragmatic motivations that we embark on the quest for metrics for intelligence of constructed systems.

---

[1] This paper is a result of collective efforts to understand the problem, and the future publication based on this paper will have multiple authors. The draft was written by A. Meystel. Initial editing was done by J. Albus, E. Messina, J. Evans, D. Fogel, and W. Hargrove. These are the authors of multiple additions to the initial draft: G. Bekey, H.–H. Bothe, B. Chandrasekaran, J. Cherniavsky, A. Clerentin, P. Davis, S.

2. Intelligent Systems (or Agents)

Intelligent systems (that are also frequently called "agents") can be introduced with different levels of detail. The simplest possible and the most general model of intelligence is just a string of six consecutively functioning elements forming a loop of closure: WORLD INTERFACE, SENSORS, PERCEPTION, WORLD MODEL, BEHAVIOR GENERATOR, and ACTUATOR. The loop of closure consisting of these six modules has a flow of *knowledge* circulating within this loop and changing its form within each of the modules. It is possible to demonstrate that if one introduces the concept of intelligent agent in this simple form, a significant degree of generality is achieved in talking about a single intelligent system as a part of the overall model of functioning. Let us try to define this loop with *knowledge* circulation in it, as a scientific entity. The subsequent description of an Intelligent Agent is relevant to our needs of analysis and design. This is the list of features characteristic for an intelligent agent.

Feature 1. Intelligence is the faculty of an agent that allows to deal with *knowledge* and to achieve the externally measurable *success* under a particular *goal*.

Feature 2. The knowledge of an agent is the collection and organization of information units. Knowledge is presumed to appear as a result of the *learning* about the objects of the external world, interconnections of the objects, and processes of changes produced by the agent within this external world. These processes are characteristic for all intelligent systems.

Feature 3. The learning process is understood as recording the *experiences* encountered by an intelligent system and deriving from these experiences a new set of rules that suggests how the intelligent system should act under particular circumstances (in a particular situation and under particular goal). **Feature 3A**. Learning provides for a successful adaptation of agent (intelligent system) to changing environments, e.g. different algorithms of new rules derivation can be utilized (i.e. algorithms of reinforcement, habituation, Hebbian association, abstraction, generalization, etc.).

Learning[2] invokes special metrics that affect the way of judging the performance and intelligence of systems with learning. In the machine learning community there is a tendency to look at three metrics: the ability to generalize, the performance level in the specific task being learned, and the speed of learning. From the point of view of evaluating intelligence, the ability to generalize seems to be the most important one. Systems can do rote learning, but without generalization, one cannot apply what has been learned to future situations. Of course, if two systems were equivalent in their ability to generalize, with the same resulting level of performance, then the one which could do this faster would be "better."

Feature 4. Experiences are understood and stored as triplets of the information units "situation→action→new situation" that allow the behavior generation module of the agent to infer what is the action that is required to improve the situation (evaluation is presumed).

Feature 5. A situation is understood and stored as a complete set of sensor inputs associated with a particular moment of time in a form that allows for processing. A situation also includes the entire situational

---

Delaroche, L. Erasmus, D. Filev, L. Fogel, W. Freeman, S. Grossberg, S. Lee, P. Lima, L. Pouchard, A. Schultz, C. Weisbin, A. Yavnai.

analysis, such as the operating goals, parameters, and hypotheses about external conditions, such as enemy locations.

Feature 6. All artifacts of learning are evaluated for their desirability according to the criteria of goodness existing in this particular agent.

Feature 7. Action of an agent results in a complete set of agent motion (or behaviors) that are developed by actuators of the agent and are sensed by the agent as changes in the external world.

Feature 8. The intelligent system (or the agent) is presumed to be equipped with the relevant sets of sensors and actuators, with the information storage, an inference system and a device for value judgment that allows for ranking both the experiences and the rules and determining their preference for the goal of the system.

One can see that no degree of sophistication is discussed in this setting. All processing is explained as inference, and various versions of inference will entail different levels of sophistication. One of the important mechanisms of inference is the mechanism of generalization: An agent is capable of inferring how to find an appropriate group of objects, how to transform it into a single object, and how to derive the rules for the generalized object from the rules that were known for its components.

So far, the described system looks very cozy and almost trivial in the very beginning of its existence. However, as the amount of experiences grows, the complexity of computations grows exponentially and the efficiency of goal-oriented functioning falls. No respectable agent would allow itself to be overburdened by growing complexity. This is why the operator of *generalization* is introduced: agents cannot afford the complexity of computations. This is the main reason for the emergence of mechanisms of generalization: they create new objects by the virtue of merging similar objects delivered and utilized by the original set of sensors and produced by actuators[3].

These generalized objects form a new world of representation: the one belonging to a lower level of resolution. As a result, we end up with a multiplicity of interrelated hierarchies of percepts, concepts, commands and actions. Corresponding multiscale systems of objects form a storage of the World Representation[4]. Any functioning actor has this system that provides its functioning.

Feature 9. The goal is the overall assignment to the system that determines the purpose of its functioning and the preferences that system uses to choose the action, and eventually determines the structure of its knowledge representation.

---

[2] Contributed by A. Schultz

[3] Generalization and abstraction occur on items resident in memory, in an indefinite amount of time. I reflect on events from last year, yesterday, and this morning, and may detect a pattern I hadn't noted before. This may be a higher-level generalization & abstraction than of the immediate kind applied to sensory inputs.

[4] We are familiar with the fact that some researchers disagree with the need for World Representation. It could be argued that all architectures are equipped by some form of World Representation, albeit under a different label.

Feature 10. In the system with a multiscale knowledge representation the action determined at the lower level of resolution becomes a goal for the higher level of resolution. Thus, we are used to the situation that the goal arrives from the exterior of each level of resolution.

Feature X. The unknown feature.

What this feature is follows from answering two questions that emerge immediately as soon as the first nine features are introduced:

Question X.1 Who creates the goal for the lowest level of resolution?

Question X.2 Can the goal be formulated internally (at a level of resolution)

The design of increasingly autonomous intelligent agents will also require an end-to-end approach, in which all the aspects of perception, cognition, emotion, and action are realized in a single system[5]. Feedback cycles of information processing need to be designed from perception through action and then back to perception again, mediated by feedback through the environment. Such cycles of information processing can evaluate the effects of system performance on the environment, and modify the system where needed to achieve better environmental control. It has also become clear that, in addition to these externally mediated cycles of information processing with the environment, internally mediated feedback is needed to achieve autonomous system properties. Such internal feedback realizes properties of intentionality and attention that are characteristic of biological intelligence.

Consciousness[6] might be considered as a possible candidate for interpreting the Feature X. This is one possible view on the contribution of consciousness as a feature (faculty) of intelligence. Only those creatures that adequately forecast their environment survive, that is, recognize the dangers and opportunities in time for a suitable reaction. Since the real world is dynamic and uncertain, having a feature for discovering new ways to solve new problems should be one of the key features of intelligence.

Consciousness provides a view of the *self* in the context of the immediate environment. As a capability it did not arise suddenly, but rather, establishes itself at different levels and in different degree. The dog understands his environment and his place within it with some degree of clarity. We know ourselves and our environment in more precise terms and can even include unseen elements. I'm conscious of the time of day, what happened yesterday, what might happen tomorrow, even what's happening in Serbia without having been there. It is consciousness that allows manipulations of alternative models of the real world as we understand them. Here is the basis for dealing with an enormous range of issues as they pertain to survival. The mechanism of consciousness seems to be the "software" of human intelligence.

The primary problem with respect to consciousness is the underlying algorithmic mechanism. This subject has received a lot of attention in recent years. The real challenge is to build a mechanism that is conscious, not simply simulates the behavior of a conscious entity. There is no homunculus within us. The question emerges, how does perception present itself to us as an integrated entity? How are we capable of understanding our own consciousness?

---

[5] These observations are taken from the abstract by S. Grossberg

[6] Contributed by L. Fogel

A related problem is concerned with "binding." In what manner are the various modalities (vision, hearing, and the other senses) combined when we now know that vision itself is compartmentalized with separate perception centers for color, shape, texture, and so forth. How can all this be done in real time? There are other intrinsic problems that are yet to be faced. An interesting question is, what will a higher level of consciousness be like, above and beyond what we now have? What if our species grows into something even more complex with greater intelligence? What would be the nature of self-awareness and understanding of the world in which it operates? Could a machine facilitate consciousness through some symbiotic relationship? There are more questions to be asked than answered. What are the links between survival and consciousness? Consciousness is essential in an n-player game wherein survival depends upon the induced behavior of other players and your relationship with them. Consciousness presumes a conscious ability. This too is an intrinsic aspect of intelligence and we expect that it shall be addressed.

3. The Problem of Measuring both Performance and Intelligence

Both engineered and organized - that is, artificially produced - intelligent systems should demonstrate qualities similar to those demonstrated by living creatures, and especially by humans: ability to work under a hierarchy of goals, and subgoals ability to perceive the external world and recognize objects, actions and situations, ability to reason, make decisions, plan, schedule and evaluate the results of actions and learn from their experiences. These systems are actually Constructed Systems with Autonomy (CSA); we will call them *Intelligent Systems*.

Intelligent Systems of interest have both their body and their mind designed by humans (engineers and programmers); we have to recognize which part of the intelligence is incorporated in their "body" and which is a faculty of their "mind" (i.e. its intelligent control system). The structure and the characteristics of the "body" can relax the requirements of the intelligent control system if the results of past experience of functioning or anticipated future situations are properly incorporated in the design. Proper distribution of systems' intelligence between body and mind is a part of engineering design. Different degrees of autonomy require different degrees of total intelligence, and a different distribution of total intelligence between the "body" and the "intelligent controller".

Intelligent Control Systems are usually equipped with a system of Perception (Sensory Processing), Knowledge Representation (where the world model is constructed, frequently in the form of the ontology), and Behavior Generation (that creates task decomposition, plans and issues commands). As a rule, these systems are multigranular (multiscale, multiresolutional), and they resolve their problems at various scales simultaneously.

Multiple existing definitions of intelligence emphasize different facets of this complex phenomenon. We will follow the definition of intelligence formulated by J. S. Albus in 1991: "…intelligence will be defined as an ability of a system to act appropriately in an uncertain environment, where appropriate action is that which increases the probability of success, and success is the achievement of behavioral subgoals that support the system's ultimate goal."[7]

Intelligent Systems differ in the depth and the breadth of the "appropriateness" of acting they demonstrate in different situations. Subsequently, they differ in the degree of "success" they are capable of achieving. The functioning could be made more appropriate and the level of success could be improved if we understand how to measure their intelligence. Thus, the measure of intelligence can be frequently reduced to measuring the "success" of functioning as provided by the ability to develop "appropriate" activities of the constructed intelligent system. The problem is non-trivial as can be seen from the case study below. We intentionally have chosen an exotic example since most of the readers can construct much more sophisticated cases related to unmanned autonomous vehicles, cooperating multilink manipulators, space stations, robot-companions, etc.

The Albus definition of *intelligence* is based upon understanding of the term success[8]. The success of solving a given task depends on the system's faculties, plus on some influences, which might be of stochastic nature or might not be measurable. One group of faculties can be called "the capacity to solve problems" or *intellect*. Intelligence includes intellect and, in addition, a number of other faculties that together help to facilitate the *success*. These additional faculties of intelligence include a) sensing abilities, b) skills of sensory processing and image interpretation, c) the capacity to collect, store and organize knowledge, d) the ability to use knowledge, i.e. via problem solving and decision making processes; the latter includes developing of the alternatives of plans for future actions, evaluating their preferability and choosing one of them, e) the ability to transform the decision into actions that lead to a success. Thus, intelligence represents a 'potential ability' to solve a given task in good time. A high intellect might compensate for the lack or deficiencies in other components of intelligence, and vice versa.

Many concepts of measuring intelligence exist. Many were proposed in communications during preparation of this White Paper. This is what L. Fogel[9] suggested:

1) Intelligence is measured in terms of the diversity of purposes that can be achieved under the range of environments. This diversity is usually reflected in the number of dimensions in the Space of Intelligence (see Section 6). The greater the diversity of purposes/situational constraints, the greater the intelligence.
2) Measures of performance must be from the point of view of some social entity. Thus, the results of measuring the degree of success are very relative. Accomplishing a certain task (or range of tasks) may be of great value to Mr. A, and of little value to Mr. B. There can be no absolute metric.

---

[7] J. Albus, "Outline for a Theory of Intelligence," *IEEE Transactions on Systems, Man, and Cybernetics,* Vol. 21, No. 3, May/June, 1991, pp. 473-509

[8] The subsequent consideration of the term *success* was proposed by H. -H. Bothe

[9] From an e-mail message, May 30, 2000

3) The worth of performance must include the cost of the performing. In some cases, this is merely operational cost, in others its R&D, T&E, acquisition and installation, as well as operations. Rarely this cost may include removal minus salvage value.

One aspect of this integrated mechanism of *intelligence* as commonly understood is that the agent who has it is often able to produce behavior that has a certain *reasonableness* to it[10]. That is, if one knew the goal of the agent, one might agree that the behavior was oriented to achieving the goal. A. Newell identified this quality of intelligence as a kind of *rationality*. He then asked what made the agent successful in achieving the goal. The answer was: the agent had knowledge and had some ways of using the knowledge for the goals.

The "way of using the knowledge" can be interpreted as and is embodied in the agent's *architecture*. He then noted that sometimes an agents's knowledge is bound up for use for only certain types of goals. On the other hand, for some agents, some of the knowledge is available for any goal for which it is potentially applicable. Chandrasekaran gives an example of a visual system that has knowledge that elements of the visual scene that have similar velocities probably belong to one object. However, while we "have" this knowledge in some sense, it is typically not available for us to reason with in our deliberative problem solving. It is simply hard-wired for use only for certain problems in vision[11].

On the other hand, we know many things explicitly. And as long as our memory doesn't fail us, we are often able to use our explicit knowledge for many different goals for which the knowledge is relevant. In the case of humans, we have a deliberative cognitive architecture that can often retrieve the relevant knowledge and make it available for the explicit (conscious) problem solving.

A. Newell proposed that an idealized version of an intelligence (in the sense of *rationality*) would always use knowledge K if it had it and if it was relevant for a goal. This is purely an architectural characterization: it doesn't say anything about what kinds of knowledge are useful. If an agent has a certain goal, if knowledge K is useful for it, and if it doesn't have it, the agent of course won't use the knowledge. But the agent probably has some other knowledge K', which may be used to generate a subgoal of identifying the knowledge needed and maybe acquiring it[12]. With the appropriate ways of interacting with the world, the agent would use knowledge K' first, and then acquire the knowledge K, and voila, the goal is achieved.

Focusing on the ability to use knowledge for any relevant goal characterized, for A. Newell, is an extremely important aspect of intelligence. We would like to notice that one more faculty of intelligence is involved: namely, focusing attention, which is frequently used by the agent in its search activities.

---

10 This discussion of the interpretation of the term intelligence was contributed by B. Chandrasekaran

11 While this thought is powerful and probably correct, the example is not particularly persuasive. It is hard to say whether this knowledge is utilized by the subject that visualizes the scene. One might assume that we group the adjacent points together into one object not because they have the same speed but, on the contrary, we deduce that they have the same speed because they belong to the same object. The grouping for declaring the fact of “belonging to the same object” might be done by the virtue of spatial adjacency no matter what the speeds of the points actually are.

12 This formidable conjecture is based upon an assumption that the agent somehow *knows* that by achieving a subgoal, knowledge about how to achieve the main goal will be acquired. Given the current state of practice, it would be more natural to assume that a problem solving intelligence should be equipped by a faculty of searching, and in situations where knowledge is lacking, it develops a set of searching activities.

This sense of intelligence goes against a common intuition in which intelligence is associated with having the knowledge rather than the ability to use the knowledge you have to acquire the knowledge (i. e. to focus attention and search). Later, we will discover that when we focus our attention and get engaged in searching, usually we end up with finding groups of similarity and create clusters[13] – objects of the lower resolution.

Newell's definition deserves our attention because it captures one sense of the term in a way to which some sort of metric may be attached. Purely reactive machines -- which map their perceptions directly into actions, such as the thermostat -- are on the low end of the scale. Further up are machines that can map from perceptions to actions by considering a large but precompiled number of alternative paths that are constructed by grouping, while groups are found by search and focusing attention.

4. A Case Study: Artificial Climate System

In an Artificial Climate System, it is required to maintain the temperature of the air in the controlled rooms within some interval of temperature $\Theta°$ (with some accuracy $\Delta\Theta°$), and provide the value of humidity within some interval of h (with some accuracy Δh) for a particular moment of time t. In addition, the Artificial Climate System should keep some function within some interval $F_t(\Theta°, \Delta\Theta°, h, \Delta h, t) \leq \Delta F$ experimentally determined to be preferable for a human being. In this case, the goal pursued by the system is not a particular state $S_t(\Theta°, \Delta\Theta°, h, \Delta h, t)$ but is rather an unknown function $F_t(\Theta°, \Delta\Theta°, h, \Delta h, t) \leq \Delta F$.

This problem is rather a nontrivial one. It can be compared with a problem of welding control where the function of the seam quality is very complicated and typically unknown since it depends on many factors, some of them hard to measure, or even evaluate. Generally, the problem is similar to the problem of optimum control of all multivariable stochastic controllers with incomplete available information that do not pursue a particular state but rather being within an interval of some cost-function. The explicit or implicit ability of a system to generalize might be crucially important for providing a proper functioning of the system and maintain the proper climate to the full satisfaction of the user.

Even more complicated functioning can be expected if this cost function is unknown, and the system of Artificial Climate should learn it by observing the behavior of the human user. This would require observing how many times the human user was turning "on" or "off" the ventilator, how many times the user was turning "on" or "off" the cooling unit, the humidifier, and what were the measures of temperature and humidity at these moments in the room. A simplistic automated system might be confused, but an intelligent system with elements of learning will pursue a mutually satisfactory schedule of functioning for all interrelated subsystems. The system will in fact learn the climate related "personalities" of the users and will learn to recognize who demands what and when. Even more bold generalizations could be expected if the system can correlate the user's behavior with the readings of temperature and humidity outside (not only inside!).

The goal of this learning process should be reduction of the amount of human intervention – that is, increasing the autonomy of the system. If the human-user needs to tune the system less frequently, this would

---

[13] One of the elements of new knowledge generation.

mean that the system works better. An even more interesting situation might happen if there is more than one user, and different users have different policies of tuning the system up, i.e. multiple users have different propensities in intervening with the Artificial Climate System. The Artificial Climate System that would minimize the total number of cases of human intervention would be considered a system for achieving consensus in a particular multi-player game.

A further development of this system might be required if the owner of this particular hotel wants actually to reduce the cost of energy required for keeping the customers satisfied. Then, the system can be designed so that it will learn habits of the customers to keep their average number of complaints below some particular level, while the energy consumed will be minimized. We can see that all these systems have a pretty high degree of autonomy: they autonomously assign the schedule of subsystems functioning. On the other hand, these systems are *subserviently autonomous*, i. e. they control their own behavior but the goals are totally determined from the external user.

The solution of this problem might be different for the systems that have a sense of *self*. A system may be considered to have a sense of self if it is equipped to take into consideration its own interests or advantage – and generate goals and success criteria for itself. Initially, we consider a set of regular obedient controllers that are intelligent (to a degree) but do not have any *self*, yet. The system equipped with a *self*, will try to keep all sources of assignment satisfied (including customers and the hotel owner) while worrying primarily about enhancement of its own life span (reducing aging, increasing reliability, and so on). In other words, a further development of the system presumes its self-evolving and self-improving.

This Artificial Climate System with elements of autonomy can be qualified as an intelligent one. It definitely should have elements of learning, should have an ability to recognize phenomena of the external world that are required for its functioning, must use elements of deductive and inductive reasoning, and must generalize upon the input information and the results of its own functioning. We can see that the "intelligence of the system" can grow, as the goal of functioning grows in its dimensionality and levels of detail. We can judge the degree of intelligence by the breadth and depth of the goals that are achieved and the performance measures that are satisfied. We are not only interested in evaluating the correspondence between the goal and performance criteria on one side and the degree of intelligence on another side. We are interested in tools that allow for the growth of intelligence and more adequate satisfaction of the assignment.

5. The System Specifications and Vector of Performance (VP)

One specific property of intelligent systems is lack of knowledge about the future conditions of functioning. The list of variables is incomplete, the intervals of future parameter changes are uncertain, the goals to be pursued can be formulated only in general. Lack of clarity in design specifications calls for design redundancy which amounts to the need for autonomously compensating for uncertain control specifications and vaguely specified contingencies.

The system requirements identify the characteristics which the Intelligent System (e. g. unmanned ground vehicle) must possess. The choice of the specific components from the Tools of Intelligence (see the subsection on that topic) mandates which of the following capabilities are included to satisfy the specific system requirements:

- to recognize objects, actions, situations
- to infer from the recognized elements of the scene
- to search for a required object within a scene
- to remember scenes and experiences
- to interpret situations
- to evaluate objects, actions, situations, and experiences
- to learn new skills from positive and negative experiences
- to generalize upon recorded similarities and acquire new concepts
- to detect an unfamiliar object, label it, and then learn about it
- to communicate with humans and other intelligent systems
- to collaborate with humans and other intelligent systems
- to interpret its own behavior
- to adapt to new environment
- to interpret behavior of other intelligent systems
- to properly generate a solution in an unexpected situation
- to perform task decomposition
- to plan and schedule in time planned activities
- to support all modes of planning/control required.

Other system requirements can be deemed pertinent to the general architectures of intelligent systems. It seems practical to construct the Vector of Performance (VP) for each of the subsystems in full correspondence with the subsystem's specifications. We always know *quantitatively* what the outputs of interest are. The set of these outputs forms the target vector VP_T. Within the space of performance there corresponds to some particular area: the zone of performance determined by the set of specifications. After testing the real i-th system or systems we receive a real vector or set of vectors {Vi} that are supposed to be compared with VP_T.

The result of this comparison is the result of measuring a concrete V_i by determining the degree of its belonging to the zone of the performance space occupied by VP_T. Note that this is not a standard single-dimensional conventional measurement when a particular unit of measurement is introduced. Rather, this is determining the membership function in a class.

The mathematics of comparison does exist. It is not frequently applied to the realistic cases because it is not frequently requested by the professionals who are responsible for the evaluation and comparison of complex systems. However, for some particular subsystems the comparison between $\{V_i\}$ and VP_T is a common practice. We refer to the area of control systems where many comparison metrics have been developed. Some additional effort would be required to apply a similar approach for more general and difficult cases but this effort is within our reach.

In the area of intelligent systems, an additional difficulty is expected linked with the fact that a concrete system is always a hierarchy of subsystems. For each particular subsystem chosen within a concrete research and/or industrial domain, the comparison between $\{V_i\}$ and VP_T is well understood. However, not much thought was given yet to the mathematics of integrating V_i and VP_T of subsystems into V_i and VP_T of the overall system. We are optimistic about development of the appropriate techniques. In many real situations, this has been done in practice. It would be appropriate to expand the experience from real situations to the general theory of (hierarchical) vector comparison since real situations affect the architectural issues in a more relevant manner.

6. Intelligence, Goals Hierarchy, and Arbitration

A device with a very low level of "intelligence," can perform its duties and achieve the goals in an excellent way within the boundaries of its "obtuseness." Yet, a very intelligent device with the ability to make powerful generalizations of the available information, capable of performing a sophisticated processing of this information and generating new concepts often cannot perform the task as well as a simple "obtuse" device, for example, maintenance of the temperature in the room within a concrete interval. This very intelligent device starts interfering with the level of humidity, looks for correlation links between recent commands of the human operator, and doing other things that the user does not need. Thus the user response: what is the merit of "intelligence" if the job has not been done or has not been performed in a timely manner (i. e. within the specified concrete interval)? Similar things happen with humans when an overeducated person is assigned for a simplistic job.

Intelligent behavior is characterized by flexible and creative pursuit of endogenously defined goals[14]. It has emerged in humans through the stages of evolution that are manifested in the brains and behaviors of other animals. Intentionality is a key concept by which to link brain dynamics to goal-directed behavior. The archetypal form of intentional behavior is an act of observation through time and space, by which information is sought for the guidance of future action. Sequences of such acts constitute the key desired property of free-roving, semi-autonomous devices capable of exploring remote environments that are inhospitable for humans. Intentionality consists of the neurodynamics by which images are created of future states as goals, of command sequences by which to act in pursuit of goals, of predicted changes in sensory input resulting from intended actions (reafference) by which to evaluate performance, and modification of the device by itself for learning from the consequences of its intended actions. Imagination images, i. e. the images of the future states produced by the planner and/or the predictor, or the results of simulation can be produced in the form in which the SP system would see if the actions were carried out, or in a symbolic form of topographical map representation (at the lower resolution), or even in a descriptive form (at the lowest level of resolution).

Intelligent Systems are to be used in cases that are too complicated for using simple controllers; otherwise simple programmed and/or automated devices should be used. A notion of *closed* vs. *open* systems should be introduced that is relevant to the situations where programmed vs. intelligent devices can be utilized. *Closed systems* can be characterized by having a clear assignment of the problem to be solved, and a crisp

---

[14] From the abstract submitted by W. Freeman

ability to be characterized by a complete list of concrete user specifications in the terms of measurable variables. These are the cases where using an intelligent system is excessive.

On the contrary, in an *open system*:

- the problem is not totally clear
- its parts are not concretized; decomposition is not obvious
- the variables are not listed in the beginning of design process
- many variables will emerge during the process of functioning
- the methods of their observations and registration are not known *a priori*
- many rules of action should be learned during the process of functioning.

So far, we can indicate two diametrically opposite strategies exercised by intelligent systems: one strategy is characterized by a very long-term general goal, say, survival of a system, another by a set of short term particular goals. The strategy of survival demands that intelligent systems be able to adapt to the environment and all circumstances. The strategy of "adapting no matter what" determines particular laws of an intelligent systems's functioning. The other strategy is "following particular goals" no matter what. The latter strategy frequently leads to the destruction of the system at hand: it might perish while following its goals persistently. Adaptation is not possible under the second alternative of intelligent systems since adaptation demands a compromise of the particular short term goals that the system was assigned.

There is an intuitive feeling that the systems with the second strategy are somehow better, or preferable than the systems that adapt no matter what. However, this intuitive feeling is difficult to rationalize and explain. Obviously, these goals belong to different levels of granularity (scale, resolution) and they can be reconciled only by considering the larger scope of the situation. Following the particular goal no matter what may lead to the destruction of this particular system but will provide for survival of the rest of the team of intelligent systems (e.g. a squad of unmanned autonomous vehicles; in other examples analogous situation takes place, i.e. as the problem gets complicated, the solution moves to the domain of multiagent solutions).

Therefore, these two strategies can be compared with respect to some additional criterion that has a higher priority than "just survival", or "just pursuing the goal." One of such external criteria is that of "knowledge acquisition." Under this criterion, one should carefully analyze the very intention to survive while abandoning the goal, or an intention to achieve the goal, even if the perspective of being destroyed is actually an imminent one. Both intentions might turn out to be secondary issues if the rate of knowledge acquisition is at stake, and in one case this rate was higher than in another. Indeed, one can adapt to the details of surrounding environment even without knowing the broader world.

In the meantime, while the system is studying the world and ardently acquires knowledge of it, the model of the world evolves so much that a *simple* adaptation is merely impossible[15], and the survival is achieved for the system that has *evolved*. Negotiation is a powerful tool that allows for adjusting the intentions (toward the goal achievement) to the rational evaluation of the losses that might occur if the goals are pursued persistently and incessantly.

The possibility of negotiation and arbitration generates more complex scenarios[16]. Assume an agent "wins" a particular negotiation at a given time. It would also allow for (but not require) that the tie-breaking arbitration assures that all of the goals are brought to the attention of negotiating parties over time. The arbitrator might want to make sure that a given agent wins "something," especially after losing out several times. Or, in goal terms, the arbitrator might be concerned about maintaining a balance "in portfolio terms". If there were goals associated with efficiency and discovery, the arbitrator might keep track of how cumulative efficiency and discovery supplement the awards of goals achievement. If, as the result of a number of decisions over time, efficiency was always winning out, and the locker of discovery items was empty, then the arbitrator could adjust his tie-breaking rules. This means that autonomous intelligent agents should not always try to be (locally) efficient, especially if they are equipped with learning.

On the other hand, we may be getting intelligences and goals of our agents mixed up. Suppose that there is a set of goals (G_1,...G_n). Different agents might have different pure goals, or they might put different weights on the various goals. Further, they might be better or poorer at pursuing those goals in differing contexts. That is, they might have different components of intelligence (I_1,I_2,..I_s) and these would be more or less important in the different contexts (C_1,...C_q). Indeed, a human may value beauty, order, material things, family, and learning new things, just to mention a few items.

This human might be very good at aesthetic matters and family matters, but not so good at order and material things. The agent might be good at trying and learning about new aesthetics-related things, but poor at doing anything risky. It is typical for humans to have a portfolio of "intelligences" as well as "goals." It would give some value to all the different goals, and would have some value to each dimension of intelligence. Another human might be characterized as an explorer, although he would value family and wealth to some degree--just not as much as new discoveries. Yet a third might be an explorer in search of tidiness (e.g. a scientist). What do you think, which human will do better? It depends. An unequivocal answer might be impossible at a single level of resolution because the true result depends on the distribution of the types of agents and the contexts that the groups of agents find themselves in.

7. What Constitutes the Vector of Intelligence (VI)?

We are still in limbo about what we should measure to evaluate intelligence: the mysterious Vector of Intelligence (VI), or the system's success as attributable to its intelligence. (The need to construct a VI emerges in many areas.)

For example, the problem of the appropriate degrees of generalization, granularity, and gradations of intelligence occurs in ontology development[17]. What constitutes the appropriate scope and levels of detail in an ontology is practically driven by the purpose of the ontology. The ability to dynamically assume one level of

---

[15] *Adaptation* is understood as a mere parametric adjustment while the evolutionary changes in the *structure* of a system are results of *learning*.

[16] Contributed by P. Davis

[17] Submitted by L. Pouchard

detail among many possible details is important for an intelligent system. It might depend on the purpose of a system. In that sense the long term purpose of the system is different from its short term or middle term goals. Clearly, the long term purpose and the multiple term goals are goals belonging to different levels of resolution and should be treated in this way. This brings us back to the measures of intelligence through success: is intelligence to be measured by the ability of a system to succeed in carrying out its goals?

The term "success" is a key word in the Albus definition, because it becomes a source of emerging gradations in intelligence, the degrees of intelligence depend on the essence of the definition of the word *success*. This means that if success is defined as producing a summary of the situation (a generalized representation of it), the summary can be computed in a very non-intelligent manner especially if one is dealing with a relatively simple situation. Indeed, in primitive cases, the user might be satisfied by composing a summary defined as a "list of the objects and relationships among them" i.e. a subset of an entity-relational (ER) network[18]. On the other hand, the summary can be produced intelligently by generalizing the list of objects and relationships to the required degree of quantitative compression with the required level[19] of the context related *coherence*. Thus, *success* characterizes *intelligence* if the notion of *success* is clearly defined.

The need in gradations of intelligence is obvious: we must understand why the probability of success increases, because somebody is supposed to provide for this increase, and somebody is supposed to pay for it. This is the primary goal of our effort in developing the metrics for intelligence. The problem is that we do not yet know the basis for these gradations and are not too active in fighting this ignorance. What are these gradations, how should they be organized, what are their parameters that should be taken into account? We can introduce parameters such that each of the parameters affects the process of problem solving and serves to characterize the faculty of intelligence at the same time.

The following list of 25 items should be considered an example of the set of coordinates for a possible Vector of Intelligence:

(a) memory temporal depth
(b) number of objects that can be stored (number of information units that can be handled)
(c) number of levels of granularity in the system of representation
(d) the vicinity of associative links taken into account during reasoning of a situation, or
(e) the density of associative links that can be measured by the average number of Entity-Relation (ER)-links related to a particular object, or
(f) the vicinity of the object in which the linkages are assigned and stored (associative depth)
(g) the diameter of associations ball (circle)

[18] See the summaries produced by the search-engines on the Web: to have it “quick and dirty” the first sentence, or the first 5 lines of an article is considered to be a summary, why not?

[19] Summarizing an article (in unstructured natural language), if done properly, is a result of generalizing the natural text description and transforming a narrative from one level of resolution by a narrative from another level of resolution.

The association depth does not necessarily work positively, to the advantage of the system. It can be detrimental for the system because if the number of associative links is excessively large the speed of problem solving can be substantially reduced. Thus, a new parameter can be introduced

(h) the ability to assign the optimum depth of associations

(This is one more example of recognition that should be performed, in this case, within the knowledge representation system).

Functioning of the behavior generation module evokes additional parameters, properties and features:

(i) the horizon of planning at each level of resolution

(j) the horizon of extrapolation at a level of resolution

(k) the response time

(This factor should not be confused with a horizon of prediction, or forecasting which should combine both planning and extrapolation of recognized tendencies).

(l) the size of the spatial scope of attention

(This corresponds to the vicinity of the associative links pertinent to the situation in the system of knowledge representation)

The following parameters of interest can be tentatively listed for the sensory processing module:

(m) the depth of details taken into account during the processes of recognition at a single level of resolution

(n) the number of levels of resolution that should be taken into account during the processes of recognition

(o) the ratio between the scales of adjacent and consecutive levels of resolution

(p) the size of the scope in the most rough scale and the minimum distinguishable unit in the most accurate (highest resolution) scale

It might happen that recognition at a single level of resolution is more efficient computationally than if several levels of resolution are involved. A finer system of *inner* multiple levels of resolution can be introduced at a particular level of resolution assigned for the overall system (e.g. Burt's pyramids[20]). The latter case is similar to the case of unnecessarily increasing the number of associative links during the organization of knowledge.

Spatio-temporal horizons in knowledge organization as well as behavior generation are supposed to be linked with spatio-temporal scopes admitted for running algorithms of generalization (e.g. clustering). Indeed, we do not cluster the whole world but only the subset of it which falls within our scope. This joint dependence of clustering on both spatial relations and the expectation of their temporal existence can lead to non-trivial results.

One should not forget that generalization (the ability to come up with a "gestalt" concept) is conducted by recognizing an object within the chaos of available spatio-temporal information, or a more general object within the multiplicity of less general ones. The system has to recognize such a representative object, event, or action if they are entities. If the scope of attention is too small, the system might not be able to recognize the entity that has boundaries beyond the scope of attention. However, if the scope is excessively large, then the system will perform a substantial and unnecessary job (of searching and tentatively grouping units of information with weak links to the units of importance).

Thus, any system should choose the value of the horizon of generalization (that is the scope of the procedure of *focusing of attention*) at each level of resolution (granularity, or scale).

All of these parameters characterize the realities of the world and the mechanisms of modeling that we apply to this world. These parameters do not affect the user's specifications of the problem to be solved in this system. The problem is usually formulated in the terms of hereditary modeling that might not coincide with the optimum modeling, or with the parameters of modeling accepted in the standard toolbox of a decision-maker.

The problem formulated by a user often presumes a particular history of the evolution of variables available for the needs of the intelligent system. Simultaneously, the user requests a particular spatio-temporal zone within which the solution of the problem is desirable. However, the input specifications often do not require a particular decomposition of the system into resolution levels and the intelligent system is free to select it in an "optimal" way. In other cases, the user comes up with an already existing decomposition of the system that appeared historically and must not be changed (like the organizational hierarchy of a company and/or an Army unit). Sometimes, it is beneficial to combine both existing realistic resolution levels and the "optimal" resolution levels implied by the optimum problem solving processes.

The discrepancy between these decompositions requires a new parameter of intelligence

(q) an ability of problem solving intelligence to adjust its multi-scale organization to the hereditary hierarchy of the system, this property can be called "a flexibility of intelligence"; this property characterizes the ability of the system focus its resources around proper domains of information.

In the list of specifications of the problem the important parameters are

(r) dimensionality of the problem (the number of variables to be taken into account)

(s) accuracy of the variables

(t) coherence of the representation constructed upon these variables

For the part of the problem related to maintenance of the symbolic system, it is important to watch the

(u) limit on the quantity of texts available for the problem solver for extracting description of the system[21]

and this is equally applicable for the cases where the problem is supposed to be solved either by a system developer, or by the intelligent system during its functioning.

---

20 P. J. Burt, “Multiresolution Techniques for Image Representation, Analysis, and ‘Smart’ Transmission,” SPIE Conference 1199: Visual Communications and Image Processing IV, Philadelphia, Nov. 1989.

21 Most of the input knowledge arrives in the form of stories about the situation. These stories are organized as a narrative and can be considered *texts*. In engineering practice, the significance of the narrative is frequently (traditionally) discarded. Problem solvers use knowledge that has been already extracted from the text. How? Typically, this issue is never addressed. Now, the existing tools of text

(v) frequency of sampling and the dimensionality of the vector of sampling

Finally, the user might have its vision of the cost-functions of his interest. This vision can be different from the vision of the problem solver. Usually, the problem solver will add to the user's cost-function of the system an additional cost-function that would characterize the time and/or complexity of computations, and eventually the cost of solving the problem. Thus, additional parameters:

(w) cost-functions (cost-functionals)

(x) constraints upon all parameters

(y) cost-function of solving the problem

This contains many structural measures. We need to trace back from an externally perceived measure of "success" or intelligence to a structural requirement. E.g, the construction codes specify thickness of structural members, but these dimensions are related to the amount of weight to support – the performance goal is the lack of building collapse.

Important properties of the Intelligent Systems are their ability to learn from the available information about the system to be analyzed. This ability is determined by the ability to recognize regularities and irregularities within the available information. Both regularities and irregularities are transformed afterwards into the new units of information. The spatio-temporal horizons of Intelligent Systems turn out to be critical for these processes of recognition and learning.

Metrics for intelligence are expected to integrate all of these parameters of intelligence in a comprehensive and quantitatively applicable form. Now, the set $\{VI_{ij}\}$ would allow us even to require a particular target vector of intelligence $\{VI_T\}$ and find the mapping $\{VI_T\} \rightarrow \{VI_{ij}\}$ and eventually, to raise an issue of design: how to construct an intelligent machine that will provide for a minimum cost (C) mapping

$$[\{VP_T\} \rightarrow \{VI_{ij}\}] \rightarrow \min C.$$

By the way, has this ever been done for the systems that are genuinely intelligent? Of course, this question is not related to design, just to measurement.

8. The Tools of Computational Intelligence

Proper testing procedures should be associated with the model of intelligence presumed in the particular case of intelligence evaluation. It seems to be meaningful to compare systems of intelligence that are equipped with similar tools. In this section we introduce the list of the tools that are known from the common industrial and research practice of running the systems with elements of autonomy and intelligence. It is also expected that these tools can be used as components of the intelligent systems architectures. Thus, they might help in developing and applying types of architectures that will be used for comparing intelligence of systems.

The following tools are known from the literature as proven theoretical and practical carriers of the properties of intelligence:

- Using Automata as a Generalized Model for Analysis, Design, and Control

---

processing allow us to address this issue systematically and with a help of the computer tools of text processing.

- Applying Multiresolutional (Multiscale, Multigranular) Approach
    1. Resolution, Scale, Granularity: Methods of Interval Mathematics
    2. Grouping: Classification, Clustering, Aggregation
    3. Focusing of Attention
    4. Combinatorial Search
    5. Generalization
    6. Instantiation
- Reducing Computational Complexity
- Dealing with Uncertainty by
    - Implanted compensation at a level (feedback controller)
    - Using Nested Fuzzy Models with multiscale error representation
- Equipping the System with Knowledge Representation
- Learning and Reasoning Upon Representation
- Using bio-neuro-morphic methodologies
- General Properties of Reasoning

—Quantitative as well as qualitative reasoning

—Generation of limited suggestions, as well as temporal reasoning

—Construction both direct and indirect chaining tautologies (inferences)

—Employing non-monotonic as well as monotonic reasoning

—Inferencing both from direct experiences as well as by analogy, and

—Utilizing both certain as well as plausible reasoning in the form of

1. Qualitative Reasoning
2. Theorem Proving
3. Temporal Reasoning
4. Nonmonotonic Reasoning
5. Probabilistic Inference
6. Possibilistic Inference
7. Analogical Inference
8. Plausible Reasoning: Abduction, Evidential Reasoning
9. Neural, Fuzzy, and Neuro-Fuzzy Inferences
10. Embedded Functions of an Agent: Comparison and Selection

Each of the tools mentioned in the list allows for a number of comprehensive embodiments by using standard or advanced software and hardware modules. Thus a possibility of constructing a language of architectural modules can be considered for future efforts in this direction.

9. The Architectures of Intelligence

Listings of all tools of computational intelligence presently available and all properties of intelligence measurable would not characterize the system exhaustively and would not suggest how to test the system. How these tools are attached to each other – this is what matters! It turns out that the architecture of the system can be decisive in providing active features of various intelligent systems.

Architectures of intelligent systems should support:

- Expected long-term mission planning (e.g. overall path planning and replanning for the whole mission performance)
- Various principles of knowledge representation
- Navigation, guidance and motion control with self-orientation using a set of techniques specified by the mission
- Auxiliary activities which require using additional intelligent control systems (e.g. for manipulator arms installed at the mobile autonomous platform)
- Ability to acquire the data, which characterize and quantitatively measure mission performance
- Perception capabilities: the character of the architecture will be strongly affected by the characteristics of all the sensors to be installed on-board of the autonomous intelligent system (for example, the unmanned ground vehicle); its intelligence will be affected by the designer's decision regarding what particular vision and other off-the-shelf perception systems are to be implemented, what is the level of human supervision[22] expected in the system (full autonomy, partial teleoperation, full teleoperation, etc.)
- Ability to handle sensing, data-processing, and decision making (including planning, navigation, guidance, and control), dealing with uncertainties, especially while operating in the uncertain environment
- Ability to respond to changes in the environment or its self-state without requiring human intervention.
- Ability to optimize performance based upon some cost-function (e.g. minimum time of task execution, minimum energy consumption, minimum final error of performance, minimum risk of being detected and/or destroyed[23])
- Multi-robot (multi-vehicle, multi-system) coordination
- Robot-supervisor interaction (in a multi-robot case this may entail robots-supervisor interaction, robots-supervisors interaction[24], etc.)

---

[22] A human supervisor will directly or indirectly assist the function of perception of the first group of unmanned ground vehicles.

[23] Often, all five of these factors are important: in this case weights must be assigned. However, some theoretical difficulties should be overcome before using this case in practice.

[24] In addition to the question: how should the interaction proceed among the members of the robotic team. One can ask a similar question about the team of human operators supervising the robotic team.

- Ability to perform a variety of tasks (e.g. in the case of unmanned vehicles, the ability to perform travel, reconnaissance operations, mine neutralizing, etc.)
- Fault-tolerant, reliable, and robust operation
- Measurable architecture performance both qualitatively and quantitatively[25]
- Extensibility for improvements and adaptation to mission specifics

Other information processing functions will probably need to be supported but those listed above most strongly affect the choice of architectural approaches. It is especially relevant in the cases where we are explicitly talking about dealing with knowledge.

The first group of these implicit architectural matters[26] includes *principles of knowledge representation* accepted in a particular intelligent system. A case could be made for semantic-based knowledge representation, including tests for completeness and consistency. Although the theories for such tests exist (e.g. Process Specification Language (PSL's) completeness and consistency can be proved within situation calculus). The breadth and scope of knowledge represented in a knowledge representation system also determines and conditions its possible re-use. Perhaps re-usable devices and software processes should be considered, since such processes potentially decrease costs of further systems. One might expect that re-usability criteria could be required for characterizing the intelligence.

Another group focuses more explicitly on ontologies that demonstrate the results of generalization within the stored linguistic information. Ontology development aims at building a machine-readable semantic layer within a (software) system. Ontologies formally express the knowledge contained in an application by providing definitions for concepts, relations and functions, as well as rules for constraining the use of the terms. Ontologies contain definitions for metadata and rules that constrain the interpretation and use of metadata. Ontologies can represent relations of inheritance, aggregation and instantiation.

Ontology development supports system interoperability by solving problems related to semantic ambiguities, and by enabling semantic communication between software agents. Software agents may refer to a common ontology to exchange messages. Actually, ontologies do not carry anything different in principle from all hierarchical constructions within the knowledge base. However, they present it in a language form, for some ontologies even in a natural language form. This opens an opportunity to communicate with large and "interdisciplinary" knowledge bases in natural language.

Providing translation mechanisms for the interoperability of applications requires that applications share a common ontology or that application concepts can be represented in a formal, declarative manner. Other benefits of ontologies include reliable system specifications, accurate data and metadata descriptions, and

---

[25] This requirement should not be confused with the functional requirement of measurability of performing a particular function, and/or the overall mission such as time of arrival, or fuel consumed, or percentage of mines neutralized. Here we are talking about performance of the architecture that should be measured in terms of performing intelligent control operations (e.g. computations per alternative of solution, goodness of solutions found, etc.).

[26] Submitted by L. Pouchard

development of common data formats for collaborative analysis. Ontologies that exist for specific tasks or domains permit knowledge sharing and re-use within the domain.

The scales and scalability criteria critical for intelligent systems are represented within ontologies, too.

10. Supervisory Control and Data Acquisition

Supervisory Control and Data Acquisition involves data collection, active communication with the user, and display. This is a group of separate subsystems (actually, several levels of the architecture) within the intelligent controller. These subsystems can be equipped by additional control loops and a separate knowledge organization system required for communication. The purposes of these subsystems are:

- to prepare information relevant to the needs of corresponding levels of control and command
- to convey this information to the user or the supervisory controller
- to conduct the dialog with the corresponding level of control and command
- to display all the information in a user friendly form e.g. use of graphics, use of previously negotiated modes of demonstration and protocol of explaining the ongoing activities
- to provide alarming, warning, notification both to other subsystems as well as for the external levels of control and command
- to provide for security by allowing different levels of control and command with different privileges.
- to facilitate printing and reporting functions, storage and display of historical data to facilitate investigation of events, investigation, and other types of analysis.

11. Tests of Machine Intelligence Contemplated in the Past

1. The Turing test, or *imitation game* was proposed by A.Turing in 1950[27]. In one version of this test a human judge interrogates a program through an interface. If the program can fool the human into believing that responses come from another human and not from a computer then the program should be considered intelligent. Clearly, in this test we don't talk about intelligence as a phenomenon but rather about an ability of pretending to be intelligent. At the present time, such an approach seems to be a naive one: it determines what *seems* to be intelligent rather than what *is* intelligent.

Nevertheless, this approach has generated a lot of literature, in particular the famous problem of Chinese room[28]. J. Searle considers the following mental experiment. A person was given a set of formal rules for manipulating Chinese hieroglyphs. This person does not speak or understand written Chinese, and he does not know the meaning of these hieroglyphs, he just can distinguish them visually[29]. The rules state that if a symbol of a certain shape is given to him, he should write down another particular hieroglyph on a piece of

---

27 A. Turing, “Computing Machinery and Intelligence”, *Mind*, Vol. 59, No. 236, October, 1950, pp. 433-460

28 J. Searle, (1980) "Minds, Brains, and Programs", *Behavioral and Brain Sciences*,

paper. The rules prescribe how the groups of hieroglyphs should correspond one to another. When a set of Chinese symbols enters from outside, the person applies the rules, writes down a set of other Chinese symbols as specified by the rules, and returns the result to the external observer. The external observer perceives the result as a grammatically correct answer in Chinese. However, the person inside does not understand Chinese. (Note that the very possibility of conducting this experiment in reality is questionable: the list of required rules would be prohibitively large if the scope of questions and required answers covers a broad domain and demands for a high degree of sophistication).

Searle believes that the person in the Chinese room does exactly what a computer would be doing if it used the same rules to engage in a grammatically correct conversation in Chinese. Both the computer and our "inside" person are engaging in "mindless" symbol manipulation. This mental experiment leads J. Searle to the following statements:

Axiom 1: Computer programs are formal (syntactic) and manipulate *symbols*.

Axiom 2: Human minds have mental contents (semantics) and manipulate *meanings*.

Axiom 3: Syntax is not translated into semantics, therefore symbol manipulation does not contain any *understanding*.

Searle's argument is intended to show that implementing a computational algorithm that is *formally* isomorphic to human thought processes cannot be sufficient to reproduce the real process of *thought*. The last decade of research in the area of intelligent systems demonstrated that this reasoning is too simplistic and is not sufficient to adequately represent even existing constructed systems with autonomy (like unmanned autonomous vehicles). Searle's schemes of analyzing processes of "thinking" are overly primitive and cannot represent existing mechanisms of sensory processing, knowledge representation and behavior generation in multiresolutional systems of motion control practiced in existing autonomous vehicles. Something more is required. Researchers that develop intelligent systems challenge Searle's argument by creating new artifacts.

2. L. Zadeh's test can be formulated as follows: a paper is presented to the intelligent system, and it is supposed to transform it into a summary[30]. The quality of the summary can be judged by the ability of the system to generalize and formulate the meaning of the paper in a sufficiently concise form. No doubt, any system that can do it should be considered intelligent. Clearly, the system should be capable of generalizing. Says L. Zadeh: "…the ability to manipulate fuzzy sets and the consequent summarizing capability constitutes one of the most important assets of the human mind as well as the fundamental characteristic that distinguishes human intelligence from the type of machine intelligence that is embodied in present-day digital computers[31]."

3. Various tests can be proposed based upon more mundane but more practical evaluations of sophistication and rationality. For example, we can check a capability of a program to generate several alternative decisions for a particular situation, and to select one of them properly; or its capabilities to analyze

---

[29] A subtle detail: distinguishing and recognizing most of the hieroglyphs is a serious intellectual problem by itself!

[30] L. A. Zadeh, from his BISC letter of 1999

[31] L. A. Zadeh, "Outline of a New Approach to the Analysis of Complex Systems and Decision Processes," *IEEE Trans. on Systems, Man and Cybernetics,* Vol. SMC-3, 1973, pp. 28-44

the experimental data related to a particular physical system, and to compute a feedforward control, and to introduce a law of feedback compensation. The key issue in the last case is the ability to use the experimental data: different experimental data require different approaches to computing feedforward control, and different laws of feedback compensation. The tradeoff "feedforward vs. feedback" is a real test of intelligence as a tool for reaching successful balance under conditions of redundancy and uncertainty.

4. A. Newell has listed properties that intelligent system must have[32]:

- recognize and make sense of a scene
- understand a sentence
- construct a correct response from the perceived situation
- form a sentence that is both comprehensible and carrying a meaning of the selected response
- represent a situation internally
- be able to do tasks that require discovering relevant knowledge.

12. Who wins the competition: the Real Intelligence or the Impostor?

Using the Turing Test to evaluate intelligence has become commonplace, although as we have already mentioned above, it does not evaluate intelligence but rather the ability of a system to *pretend being intelligent*. Competitions are one of the straightforward primitive methods of judging the degree of intelligence. The deficiencies of competition are clear from the following list:

- in a competition, a random set of circumstances can affect the results rather than a set of capabilities of the competing systems; thus, only the results of multiple competitions can be valid
- it is difficult, if not impossible, to separate the part of intelligence endowed in the body design from the part of intelligence incorporated into the system of intelligent control; thus, for judging the intelligent control system, identical bodies are presumed
- competition in the natural environment cannot guarantee the equality of the problems to be encountered by competing parties; in constructed (artificial) environments, the difficulty of the problem drops drastically; it does not require that much "intelligence"

The latter feature is not necessarily always the case. The actual challenge is to provide a rich enough environment within which the tests can be conducted. An example of this would be a completely instrumented test course for evaluating autonomous robot mobility and mapping abilities, rather than the simple "box world" that is frequently used. In fact, one of the keys to our efforts in performance metrics is to come up with these sufficiently rich environments (test courses or very detailed, ground-truth simulation environments) which can be used to evaluate the performance of different systems. It is not an easy task. We should encourage a broad discussion on defining requirements for such environments.

---

[32] Newell, A. (1982). The knowledge level. *Artificial Intelligence*. 18(1), 87-127; Newell, A. and Simon, H. (1963), “GPS: A program that simulates human thought,” In *Computers and Thought*, ed. Feigenbaum and Feldman. McGraw-Hill, New York.

Therefore, winning a competition, however exciting it might be, leads to the old pitfall of the Turing test: winning requires no more than pretending to be intelligent rather than demonstrating real tools of intelligence. Testing of intelligence is a must, but the way of testing is a matter of discussion. The challenge for competitions is to overcome these obstacles. For example, developing an artificial environment that is dynamic and challenging , yet reproducible.

13. Measuring the Intelligence Contemplated for the Future

Measuring Intellifactors. One can start analyzing the problem of measuring intelligence within the domain of Albus's definition that assigns this faculty for control purposes[33]. The factors of intelligence are the factors of processes that contribute to intelligence (intellifactors). Logistically, they are dimensions of VI, mathematically, we can express this as follows:

$\mathbf{X_i} = \{\mathbf{x} | \mathbf{x}$ is all possible intellifactors$\}$

and the set of intellifactors $\mathbf{\{X_{if}\}}$, is an element of the power set of $\mathbf{X_i}$.

A measure of intelligence (IQ) is the measure that can assign a real number to the collective performance of each element in the set X_i. The measure of intellifactor (IFQ) is a measure that assigns a real value to the collective performance of each element in X_{if}.

Measuring the Power of Generalization. There exists a way to narrow the gap between building an intelligent machine (with its ontogeny[34]) and understanding the intelligence process by itself (with its epistemology[35]). The way is to model the process in a biological system[36]. How do brains do that? Brains avoid catastrophic failure when the complexity of computations grows exponentially by use of the NN-dynamics for generalization by creating "objects" (classes). It is experimentally confirmed that for the same operation of generalization, computer elements need more computations than brain needs. One can judge on the comparative productivity of computers during simple maps generalization[37] and instantaneous gestalt insights performed by the brain during human processing of complex images.

Measuring the System's Intelligence by the Degree of Uncertainty. The latter observation is linked with the entropy based considerations. Any measure of uncertainty (entropy in particular) is an acceptable measure of intelligence. If one can measure our uncertainty in taking decisions among alternatives, one can reduce this value of uncertainty (e.g., by learning), so our system is intelligent. But how do we measure the value of each alternative? Again, by its uncertainty. A possible way is to measure the probability of success of

---

33 This concept of measuring intelligence was contributed by Louwrence Erasmus.

34 or how it is done in a living organism

35 or how it is done in the theory of knowledge

36 This concept was proposed by W. Freeman. He refers to the A. Meystel’s statement "the mechanism of generalization to emerge: it creates new objects" quoted from his e-mail letters to Advisory Board Members.

37 J. D. McMahill, Interactive Generalization: User’s Guide, CMU, Pittsburgh, PA 1998; G. L. Bundy, C. B. Jones, E. Furse, “Holistic Generalization of Large Scale Cartographic Data,” in J. Muller, J. Lagrange, R. Weibel (eds.), *GIS and Generalization Methodology and Practice*, Taylor and Francis, London, 1995, pp. 106-119

meeting the specifications for each of the alternatives by successive applications or using a model (we might call it Reliability, in this sense). The higher the success, the lower the uncertainty/entropy. We may counter-balance this with the cost (or complexity) of achieving very successful alternatives (typically, the higher the reliability, the higher the cost)[38].

Constructing the Benchmarks. Judgment of the system's intelligence can be done by using indirect, albeit easy to measure values. In constructing benchmarks, we use the fact that the fundamental attributes of intelligence include:

- Ability to perform tasks in unstructured environments
- Ability to learn from experience
- Ability to transfer knowledge from one domain to another
- Ability to solve complex problems, requiring deductive and inductive reasoning

The following simple measures can be used as metrics for such abilities in machines[39]:

- Size and complexity of programs required
- Memory requirement
- Solution time

Clearly, such measures are useful only if (a) they are applied to benchmark problems, (b) all contestants use the same type and model of computer, and (c) all programs are written by comparably competent programmers, so that the programs are optimal in some sense.

Given these constraints, we could test intelligent systems A and B on the same benchmarks. The one that accomplishes the task more quickly, and does so with the least complex programs and least memory will be declared "more intelligent". While evaluating the level of intelligence based on this definition (to avoid the confusion of introducing a new one) we have to take into account[40]:

- type of uncertain environment
- strategy of achieving the goals
- capability of the system to automatically create and update its subgoals.

Most of the well-established methods for robust control design provide the capability to deal with small parametric and structural uncertainties and therefore represent a basic level of intelligence in the control system according to the definition of Albus. Situational uncertainty, e.g. drastic changes in the environment that are due to completely different operating conditions, severe and unpredictable disturbances, etc., completely alter system dynamics, and therefore require control systems with a much higher level of intelligence.

---

[38] The latter considerations were suggested by P. A. Lima
[39] Contributed by G. Bekey
[40] From the abstract submitted by D. Filev

Measuring Autonomy vs. Intelligence. The following question can be considered a fundamental one[41]: What is more important and meaningful to define and to measure with respect to the context of Intelligent Autonomous Constructed System– Autonomy or Intelligence of a Constructed System? We are looking for Autonomy, as the premier requirement of an Intelligent Autonomous System. From the designer or the user point of view, Intelligence enables Autonomy, but it is not a system design objective or a system requirement *per se*.

The definition of Autonomy is probably more precisely measurable and more meaningful and it is easier to come to a consensus about what Autonomy or an Autonomous System is all about, rather than what is Intelligence or an Intelligent System.

14. Simulated Functioning and Scaled Hardware Testing of Intelligent Systems

The hope is for a balanced combination of a) thorough simulation and b) scaled hardware testing. Many researchers focus upon simulating systems with high autonomy[42], like B. Zeigler in USA, K.-H. Brassel in Germany, I. Peters in Switzerland, J.-H. Kim and T.-G. Kim in Korea, and others. However, the challenge of evaluating intelligence of these systems remains an active problem to be resolved in the upcoming decade.

The most intricate problems associated with the variability and combinatorics of realistic situations can be resolved by simulating these situations. Thus even the predicament of absent hardware can be avoided by simulating the problem-impregnated situations. Contests and competitions can be considered a part of this paradigm. One cannot come even anywhere near covering in realistic testing the spectrum of philosophical[43] views of intelligence (just start to read the mind/body literature!) On the other hand, one might be inclined to scale back the possible analogies to human intelligence and human involved testing to less convenient but more pragmatic scenarios.

The Paradigm of Contest and Competitions

1. Symbolic systems. The a-y classification of measurable characteristics (see Section 7) can be made very representative but is definitely too constrained by the existing general systems and ways of representing information. Indeed, each of the 25 items on this list is a strong reduction of actual possibilities. Start with (a) memory temporal depth: why it should be limited? or why should only one value of depth be considered? The next item is (b) number of objects that can be stored: why should this number be limited? Then, we come to the number of levels of granularity, definitely a limitation that should depend on the problem. Then, we face limitation on the vicinity of associative links – the latter should not be limited as well! All 25 items on this list limit the opportunity to find better (not to speak about "the best") solutions. In the meantime, the environment

---

41 From A. Yavnai's abstract

42 See in Ed. by H. Sarjoughian, F. Cellier, M. Marefat, J. Rozenblit., *2000 AI, Simulation and Planning in High Autonomy Systems*, Proc. of the SCS Conference in Tucson, AZ, March, 2000

43 From the e-mail letters by J. Cherniavsky

conducive for contests and competitions is by definition oriented toward a permanent atmosphere of inventiveness and development of new signs and new phenomena to be encoded by these signs.

2. Systems with learning. Learning is never forgotten as a very important subsystem of intelligence. However, there is not too much discussion related to the nature of learning as a substitute for real contests and competitions. In the meantime, learning plays the role of rehearsing expected ("would be") situations of contest and competition. Learning via prior experiences or via planning is a mechanism that prepares a system for contingencies. Thus, learning serves as a critical characteristic of intelligence that solely determines both the success and failure. It's there, but it's primarily implicit and serves as a supportive system that serves rather for improving functioning. Learning provides for a successful adaptation of the intelligent system to changing environments, e.g. different algorithms for deriving new rules can be utilized for different cases (i.e. algorithms of reinforcement, habituation, Hebbian association, abstraction, generalization, etc.). A multiplicity of situations can be anticipated where, without learning, the central purpose of the system could not be achieved.

3. Application Focused Intelligence. In many cases, the intelligence might be defined relative to a domain of application. Even in the human cases there are people who are "car intelligent" but "literature ignorant" - different domains, different abilities. This generates a question: if in the human domain one might distinguish different types of intelligence (Gardner's 7, Sternberg's 3, etc.) – should it be beneficial to try something similar in the autonomous unmanned, or partially manned systems? Indeed, for a human, the need to quickly move from one subject-oriented vocabulary to another might create a need to deal with using domain oriented algorithms of generalization, or pattern recognition. Can it be beneficial in the unmanned cases?

All three of these questions can be resolved within the domain of contests and competitions. We can create and focus on a specific domain where things like self-sustained, appropriate behavior, ability to quickly act in an uncertain environment, etc. can be physically quantified by realistic measures of performance (units of time, money, energy). The various contests (AAAI urban search and rescue, robotic soccer, the data-mining contests, the information retrieval competitions, the speech understanding rallies, etc.) provide the plausible level to measure and thus compare systems.

15. The Intelligence of Sensing and Sensory Processing

Available results have already suggested that the brain designs for sensory and cognitive processes differ from, and are even computationally complementary to, the designs for spatial navigation and action. This complementarity can be noticed by observing that cognitive knowledge needs to accumulate in a stable way over a period of years, with new knowledge not accidentally erasing previously learned, but still useful, knowledge[44].

The problem of data fusion (both heterogeneous or homogeneous) generated a demand that the robustness of the fusion stage be closely linked to the number of significant criteria permitting to associate information required for interpretation[45]. Both the uncertainty and the error of the input data, as well as

---

[44] From the abstract by S. Grossberg

[45] Contributed by A. Clerentin and L. Delahoche

uncertainties and errors of the available internal knowledge, jointly produce the uncertainty and the error of interpretation. The uncertainty is meant to characterize the "degree of actual existence" of the data; the error characterizes imprecision on the numerical evaluation of the data. The uncertainty and error estimation in classical fusion processes are generally based on a probabilistic approach. As the number of factors to be associated for interpretation grows, the need to work with multi-criteria techniques grows. The latter should help to evaluate the performance of each stage of global fusion processing: for example, data fusion for localization (generally allows for heterogeneous fusion) or data fusion for incremental map building (generally demands for homogeneous fusion: the same kind of primitives must merge on different acquisitions). Here again the use of tools like Dempster-Shafer theory of evidence might be promising.

In a number of applications, including the area of autonomous robotics, the problem of multi-sensor fusion and joint interpretation determines the value of intelligence related to sensing and sensory processing. It is clear that, in many situations, the use of multiple sensors is the only way of dealing with the richness of the external world. Any given sensor takes information about only one of the many attributes of the environment. But often the arriving information must be carefully gleaned for more than one attribute simultaneously. Only in this case can the required depth of interpretation be achieved.

So, the problem is how to integrate the information, especially when the sensors are disparate and when the viewpoints and even scales of incoming information are different. To overcome these problems, several fusion methods are used. The majority use a probabilistic approach (Bayes rules). A significant portion use a possibilistic approach that considers sensor evidence to be the value of belief (these rely on Dempster-Shafer theory). This theory is appropriately expressive, it explicitly represents ignorance, enabling the robot to differentiate between ambiguous sensing results and not having sensed at all. Other approaches include fuzzy logic or neural networks.

Information fusion is a growing research domain and of the numerous developed applications show that it enhances the level of autonomy and intelligence of engineered systems, especially autonomous robots.

16. Questions To Be Answered

This is the list of questions that the Workshop will try to answer[46]:

Question 1. What is the vector of intelligence (VI) that should be measured and possibly used as a metric for systems comparison?

Question 2. Should VI be measured in addition to, or instead of, measuring the vector of performance (VP) determined by the standard specifications?

Question 3. If two systems have the same VP, what is implied by the difference in their VI values? Can this difference be represented in monetary (cost) units?

Question 4. Is it possible (and meaningful) to have different VI measures: a) goal-invariant, b) resource-invariant, c) time-invariant?

---

[46] Questions 4, 6, 7, 8 were contributed by S. Lee

Question 5. What should be recommended as a test of VI and how can VP be normalized so that comparisons may be performed at the same normalized value of VP?

Question 6. Does a universal measure of system intelligence exist such that the intelligence of a system can be compared independently of the given goals[47]? A goal-independent measure may be more difficult to define. A goal-dependent measure, however abstract the goal may be, can allow for a clear comparison among the systems of different architecture but with the same goal. For instance, for the latter case, an intelligence can be represented as how efficiently, and how optimally a system reaches the given goal by itself, i.e., the power of automatically solving problems defined as the discrepancy between the goal and the current state.

Question 7. Should the intelligence measure of a system be solely based on problem-solving capability at time "t" or should it contain the potential increase of problem-solving capability in the future based on learning?

Question 8. Should the resources required for building systems and system operation play a role in defining the measure of intelligence? As mentioned above, the efficiency in problem solving should be included in the measure: for instance, the time and energy required to reach a solution should be taken into consideration together with the optimality of the solution. But, it is not clear whether we should or should not include the cost of building a system.

As a reminder, a set of other questions that are ingrained (directly, or indirectly) in the main questions is formulated as follows:

Question 9. These are the less profound ("secondary") questions that should be addressed at the workshop and possibly unequivocally answered:

a) how to form VI for various architectures?

b) should the questions 1 through 5 be related to intelligent systems, or autonomous systems, or both?

c) what is the protocol for dealing with uncertainty when the uncertainty metric is to be applied in the procedures of decision making? for example, how does the uncertainty of planning affect the cost of goal achievement?

d) what are the guidelines in constructing the world model and determining its scope in the variety of applications? how does the scope of "world model" affect the sophistication of intelligent behavior?

e) how are the questions 1 through 5 related to (and the answers applied to) the systems that are working under a hierarchy of goals?

f) should a competition between intelligent systems be considered a valid method of judging VI value?

---

47 This seems to be hard to achieve for biological systems. This will be eventually addressed, but in the short term run the concrete goal of particular cases seems to be more attainable. A single measure of intelligence requires constructing a system of meta-knowledge.

17. Glossary

Autonomy – an ability to generate one's own purposes without any instruction from outside *(L. Fogel).*

Alternative definitions:

a) independence.

b) Self-government or the right of self-government; self determination.

c) Self-government with respect to local or internal affairs (AHD) ;

d) the right of self-government,

e) self-directing freedom (Merriam-Webster)

Autonomous System – a constructed system is autonomous if there is a likelihood that circumstances will arise in which no-one can predict in advance what it will do. *(T. Whalen)*

Autonomous Intelligent System - an autonomous constructed system is intelligent if we can be reasonably confident that whatever unpredictable thing it does do will be something that tends toward success in the goals for which the system was constructed in the first place. *(T. Whalen)*

Agent (sometimes **Autonomous, Intelligent**) – a term that has been introduced to use the word system which is regarded by many as a less desirable one when the software is involved, especially the one with properties of intelligence. The term Agent has some anthropomorphic overtones, Agent is presumed to be a system that probably can sense, reason and is intended to act. In other words, Agent should be understood as a system with elements of intelligence and autonomy.

Intelligence - an ability of a system to act appropriately in an uncertain environment, where appropriate action is that which increases the probability of success, and success is the achievement of behavioral subgoals that support the system's ultimate goal *(J. Albus)*

Alternative definitions:

- the ability to solve new problems in new ways *(L. Fogel)*
- the capacity to acquire and apply knowledge (AHD).
- the faculty of thought and reason (AHD).
- the ability to adapt effectively to the environment, either by making a change in oneself or by changing the environment or finding a new one (Britannica).
- the ability to learn or understand or to deal with new or trying situations (MWD)
- the skilled use of reason (MWD)
- the ability to apply knowledge to manipulate one's environment or to think abstractly as measured by objective criteria (MWD)

18. Appendix

How is Testing of Intelligence Performed on Humans?

The most widely used intelligence tests include the Stanford-Binet (SB) Intelligence Scale and the Wechsler Scales (WS). The Stanford-Binet test was first introduced in 1916 by Lewis Terman from Stanford University. The individually administered test, revised in 1937, 1960, and 1972, evaluates persons two years of age and older. It consists of an age-graded series of problems whose solution involves arithmetical, memory, and vocabulary skills. WS-test gives both the overall IQ as well as separate IQs for verbal and performance subtests. An example of a verbal subtest would be vocabulary breadth, while an example of a performance subtest would be picture arrangement, so that they tell a comprehensible story.

IQ was originally computed as the ratio of mental age to chronological (physical) age, multiplied by 100. If a child of 10 performs the test at the level of an average 12-year-old, this 10-year-old is considered to have a mental age of 12. In this case the child was assigned an IQ of (12/10)x100, or 120. The concept of mental age is not a persuasive one, and the computation of mental ages is not used frequently. The values of IQ are more persuasive if they are computed on the basis of statistical distributions.

Intelligence tests created a controversy about what kinds of mental abilities constitute intelligence and whether the IQ adequately represents these abilities. It turned out that intelligence tests give better results for rich kids and are worse for less privileged racial, ethnic, or social groups. Consequently, psychologists have attempted to develop culture-free tests that would more accurately reflect an individual's native ability. Johns Hopkins Perceptual Test, developed in the early 1960s for measuring the intelligence of preschool children, has a child try to match random forms (geometric forms, e.g. circles, squares, etc. are avoided because some children may be more familiar with them). Another solution was to use test materials pertinent to a child's living environment.

Psychometric tests are performed by observing and evaluating the performance of the Elementary Cognitive Tasks (ECTs) with items of ECT based on past acquired knowledge, reasoning, and problem solving requiring the concerted action of a number of relatively complex cognitive processes. A particular ECT is intended to measure a few relatively simple cognitive processes, independent of specific knowledge or information content.

Each ECT is devised to address a different set of cognitive processes, and performance on two or more different ECTs yields data from which individual differences in distinct processes can be measured, such as stimulus apprehension, discrimination, choice, visual search, scanning of short term memory (STM), and retrieval of information from long term memory (LTM). ECTs typically do not depend on previously learned information content, and in those that do, the content is so familiar that it should be common to all individuals undergoing the test.

Most ECTs are so simple that every tested individual can perform them easily. The differences in performance are measured in terms of response time (RT). The most interesting ECTs are those with RTs of less than one second and with response error rates close to zero. The subject's median RT (over n number of trials) and the subject's intraindividual variability of RTs (measured as the standard deviation of RT over n trials) are of particular interest. Another type of ECT, known as Inspection Time (IT), measures sheer speed of perceptual discrim ination (visual or auditory) independently of RT.

Measures of RT and IT derived from the various ECTs are analyzed and their correlation is estimated. For single ECTs, the correlations depend on the complexity or number of distinct processes involved in the ECT. Some processes are more strongly correlated than others. Interpretation of these correlations depends on the goal of testing and properties of intelligence that are tested.

A similar approach to testing particular skills can be exercised in the area of intelligent systems. Our ability to construct metrics should depend on the particular tools or facets of intelligence we will analyze as related to the particular performance results.

However, all psychological tests of intelligence have one feature in common: they rely upon successful performance of particular tasks, but they do not attempt to introduce any relatively comprehensive form of the model of intelligence. It is understandable for measuring intelligence of such an object as a human being. It would be unforgivable to impose similar detriment upon a researcher in cases where intelligent systems are autonomous mobile vehicles, organizational systems, large computer based control systems like unmanned power plants, structures of company management, stock market. If we succeed with these types of intelligent systems, we might be encouraged to attribute some model to a human intelligence.